PATENT ABSTRACTS OF JAPAN

(11)Publication number : **11-015604**
(43)Date of publication of application : **22.01.1999**

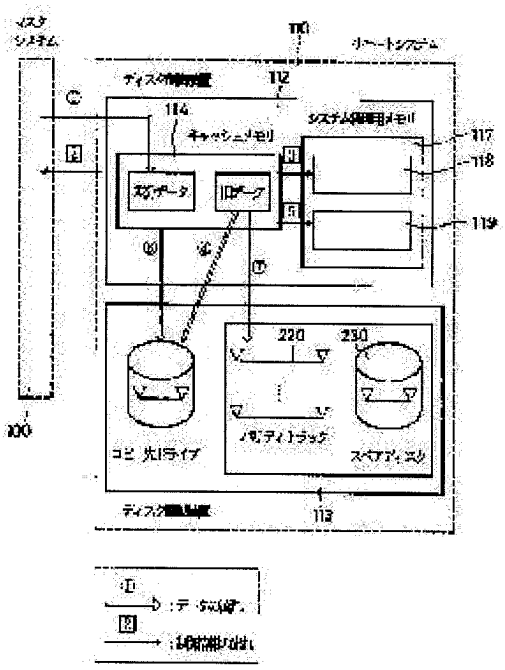
(51)Int.Cl. **G06F 3/06**
G06F 12/16

(21)Application number : **09-168898**
(22)Date of filing : **25.06.1997**

(71)Applicant : **HITACHI LTD**
(72)Inventor : **KIMURA YUKIHISA**
NAGASAWA MITSUO
NAKAMURA KATSUNORI

(54) DATA MULTIPLEX METHOD

(57)Abstract:
PROBLEM TO BE SOLVED: To improve the fault resistance of restoration copying for restoring multiplex in a data multiplex system and to maintain the matching of copy data.
SOLUTION: In restoration copying for maintaining data multiplex between a master system 100 and a remote system 110, former data on a disk drive unit 113, which corresponds to differential data, is saved to a cache memory 114 or a parity track 220 or a spare disk 230 before overwriting and differential data are stored in the target storage area of the disk drive unit 113 on the side of the remote system 110 receiving differential data from the master system 100. When restoration copying is failed owing to the fault of a master system 100-side, former saved data are returned to the disk drive unit 113 and the matching of copy data is maintained.



LEGAL STATUS

[Date of request for examination]	16.04.2001
[Date of sending the examiner's decision of rejection]	27.07.2004
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	
[Date of registration]	
[Number of appeal against examiner's decision of	

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（Ａ）**　(11)特許出願公開番号

特開平11－15604

(43)公開日　平成11年(1999) 1月22日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｇ０６Ｆ　3/06	３０４	Ｃ０６Ｆ　3/06	３０４Ｅ
12/16	３１０	12/16	３１０Ｊ

審査請求　未請求　請求項の数３　ＯＬ　（全 20 頁）

(21)出願番号　特願平9－168898

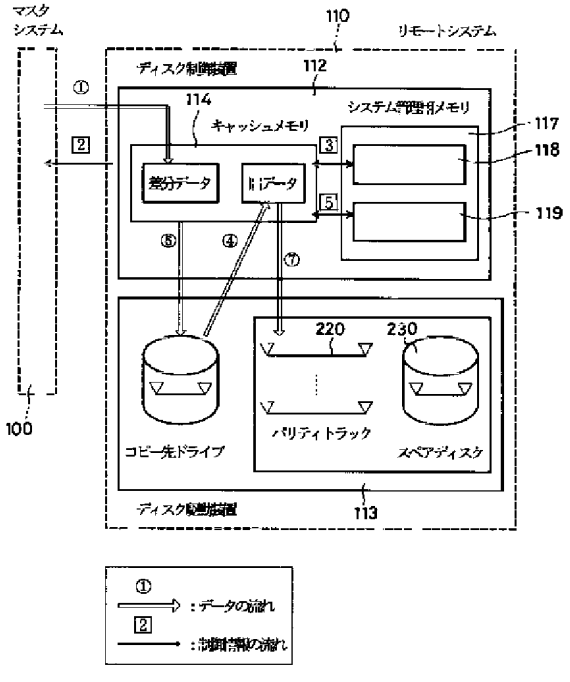
(22)出願日　平成９年(1997) ６月25日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(72)発明者　木村　恭久
神奈川県小田原市国府津2880番地　株式会社日立製作所ストレージシステム事業部内
(72)発明者　長澤　光男
神奈川県小田原市国府津2880番地　株式会社日立製作所ストレージシステム事業部内
(72)発明者　中村　勝憲
神奈川県小田原市国府津2880番地　株式会社日立製作所ストレージシステム事業部内
(74)代理人　弁理士　筒井　大和

(54)【発明の名称】　データ多重化方法

(57)【要約】
【課題】　データ多重化システムにおける多重化回復のための回復コピーの障害耐性を向上させ、コピーデータの整合性を維持する。
【解決手段】　マスタシステム１００とリモートシステム１１０との間におけるデータ多重化を維持するための回復コピーにおいて、マスタシステム１００から差分データを受け取るリモートシステム１１０の側では、上書きする前に、当該差分データに対応したディスク駆動装置１１３上の旧データを、キャッシュメモリ１１４やパリティトラック２２０あるいはスペアディスク２３０に退避させた後、当該差分データをディスク駆動装置１１３の目的の記憶領域に格納する。マスタシステム１００側の障害等で回復コピーが失敗した場合には、退避してあった旧データをディスク駆動装置１１３に戻して、コピーデータの整合性を維持する。

図8



【特許請求の範囲】
【請求項１】　第１のシステムに設けられた第１の記憶装置に格納されたデータを、少なくとも一つの第２のシステムに設けられた第２の記憶装置に複写することで前記データの多重化を行うデータ多重化方法であって、
任意の要因にて前記第１のシステム側に蓄積された未複写の前記データを前記第２のシステム側に転送する回復複写処理の実行に際して、前記第２のシステム側では、前記第１のシステムから到来する前記データに対応した旧データを任意の方法にて保存しておき、前記回復複写処理が失敗した時には、保存されている前記旧データを用いて前記第１のシステムと前記第２のシステムとの間における前記データの整合性を維持することを特徴とするデータ多重化方法。
【請求項２】　請求項１記載のデータ多重化方法において、
前記第２のシステムでは、
前記第１のシステムから到来する前記データに対応した前記旧データを前記第２の記憶装置における現在の記憶領域から他の任意の記憶領域または他の任意の記憶媒体に退避させることで前記旧データの保存を行う第１の操作、
前記第１のシステムから到来する前記データを、当該データに対応する前記旧データの格納領域とは異なる格納領域または記憶媒体に一時的に保持することで前記旧データの保存を行う第２の操作、
前記回復複写処理の完了までの許容最大時間が任意に設定され、前記許容最大時間を超過しても前記第１のシステム側から前記回復複写処理の完了報告がない場合には、前記回復複写処理が失敗したものと見なして前記旧データによる前記整合性の維持を行う第３の操作、
の少なくとも一つの操作を行うことを特徴とするデータ多重化方法。
【請求項３】　請求項１または２記載のデータ多重化方法において、
前記第１のシステムでは、任意の前記要因によって、当該第１のシステム内で発生した前記データを前記第２のシステムに複写できずに蓄積するとき、更新が発生した順序が弁別可能に前記データを蓄積し、前記回復複写処理の実行時には、最も過去に発生した前記データから順に時系列に前記第２のシステムに転送する第４の操作を行うことを特徴とするデータ多重化方法。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、データ多重化技術に関し、特に、例えば遠隔地に設置された複数のシステム間にてデータを複写することでデータ多重化を実現する場合におけるデータの整合性の維持管理等に適用して有効な技術に関する。
【０００２】
【従来の技術】情報処理システムは、社会活動の広汎な領域に利用されており、その故障は重大な社会的混乱を招く要因となり得る。このため、例えば、取り扱うデータの喪失が許されない金融等の分野を初めとして、以前より、取扱うデータを失うことのないよう、運用システム配下において、常時バックアップデータを採取保存することは基本的に行われてきた。
【０００３】ところが、最近の天災や事故の教訓から、通常運用しているマスタシステムが復旧不可能な場合に陥り、大量のバックアップデータやログからのデータ修復に要する手間や時間、更には情報を失うことなどの問題点を考慮し、遠隔地に設置されているリモートシステムにバックアップコピーデータを貯えるための方法が、例えば特表平８－５０９５６５号公報に開示された「遠隔データのミラー化」等の技術として新たに提唱された。
【０００４】その方法は、マスタシステムのデータをそのままリモートシステムに反映することにより「ミラー状態」を維持するものであり、マスタシステムでの運用が不可能となった場合に、リモートシステムに運用を移行し、システム運用再開を一段と容易にしようとするものである。
【０００５】また、リモートシステムへのバックアップコピー方法としては、マスタシステムとリモートシステム間のデータの更新契機から、大きく「同期型」と「非同期型」の２種類に分けることができる。
【０００６】前者は、マスタシステムのホストより発生した各更新Ｉ／Ｏ命令に対し、まずマスタサイドの記憶装置に書き込みを行い、続いてリモートシステムの記憶装置に向けて書き込みを実施し、リモートシステムより書き込み終了通知を受領したところで、マスタサイドの記憶装置はマスタシステムのホストに対して、最終的な書き込み終了報告を行い、常にマスタサイドとリモートサイドの更新は同期を保つ技術である。
【０００７】それに対し、後者は、マスタシステムのホストより発生した各更新Ｉ／Ｏに対し、マスタサイドの記憶装置に書き込みを終了した時点で、マスタシステムのホストに書き込み終了報告を行い、この更新Ｉ／Ｏに対するリモートシステムへの更新は遅れて、即ち、非同期に実施される技術である。
【０００８】通常の「ミラー状態」を維持している条件にあって、仮にマスタシステムが重大災害等で運用不可能に陥った場合、「同期型」においては、データ更新の同期性から「ミラー状態」は維持される。この場合、最終更新に関するデータ分は、タイミングによりリモートシステムに反映されないかもしれないが、整合性に関しては問題ない。一方、「非同期型」においては、リモートシステムには非同期に更新データが反映されるため、場合によっては「ミラー状態」どころか整合性の保証もできなくなる。

【０００９】この整合性の維持の方法としては、例えば、特開平６－２９０１２５号公報に開示されているように、更新データを待ち行列化して、更新があった順にリモートシステム側のバックアップ記憶媒体上に反映を行う技術が挙げられる。

【００１０】

【発明が解決しようとする課題】上述のような従来の技術でのリモートシステムへのバックアップコピー方法は、常にリモートシステムが正常に動作しており、常時バックアップコピーが可能な条件が成立する前提でのみ、リモートバックアップコピーとしての意味をもつものである。

【００１１】具体的に例を挙げて説明すると、リモートシステムへのバックアップ運用を実施している最中に、マスタシステムとリモートシステムを結ぶデータ転送路において自動的あるいは、センタ運用者や保守員等による人手介入によって修復可能なレベルの一時的障害（簡単な例を挙げれば、データ転送路としてのケーブルがはずれたケースなど）が発生した場合、その一時的障害が取り除かれるまでの期間は、リモートシステムへのデータの反映が不可能となる（この状態を「サスペンド状態」と呼ぶ）。

【００１２】もちろん、この「サスペンド状態」は短期的な場合もあれば、数時間以上を費やす長期的な場合もあり得る。

【００１３】一方、マスタシステムにおいては「サスペンド状態」中もデータの更新は、絶え間なく次々と行われる（この「サスペンド状態」中に発生する各更新データを「差分データ」と呼び、総称して「差分データ群」と呼ぶ）。

【００１４】その後、「サスペンド状態」が解除され、バックアップコピーが再開可能なノーマルな状態に戻ると、マスタシステムはリモートシステムに向けて、「差分データ」の反映を行いつつ、その間においても逐次入る更新データに対する反映も考慮にいれ、データの「ミラー状態」への回復の為の試行（本試行を「回復コピー」と呼ぶ）を実施する。

【００１５】ところが、この「回復コピー」は、バックアップコピーとしての意味を失う程の重大な問題を抱えている。

【００１６】一般に、マスタシステム側では次々と入ってくる「差分データ」をマスタシステム側の記憶媒体上に書き込み、記憶媒体領域全体をビットマップ形式で表現し、更新のあった領域に対応するビットを立てることにより、「差分データ」の格納領域を表現した「差分ビットマップ」にて管理している。ところが、複数の「差分データ」：データ１，２，３，．．．が存在した場合、「差分ビットマップ」では、更新の行われた順番まではわからない。そのため、「回復コピー」が始まると、「差分ビットマップ」を最初から検索し、差分のある領域をみつけて、そのデータをリモートシステムに反映している。そのため、「回復コピー」の最中にマスタシステムが重大障害（自然災害、人為災害）等によって、データの読み出しが出来ないなどの運用不可状態に陥った場合、リモートシステムでは、運用上情報の整合性を保証する必要のある単位（それは、１つのデータセット単位であったり、複数のお互いに関連のあるデータセットで構成される単位であったりする。その他、物理的あるいは論理的に定義されている各デバイスも１つの単位として考えられる。）での整合性の保証をすることができないため、この単位でのバックアップコピーは運用出来ず、マスタシステムの破壊状態によっては最悪の場合、データ喪失の状態に陥る。

【００１７】本発明の目的は、複数のシステム間でのデータ多重化を維持するための差分データの回復複写における障害耐性を向上させることにある。

【００１８】本発明の他の目的は、マスタシステムおよびリモートシステムが、ある整合性の維持を要求する情報単位において、特に「サスペンド状態」に陥る瞬間まで「ミラー状態」を維持していた場合の、「回復コピー」におけるリモートシステム側での多重化データの整合性を保証する技術を提供することにある。

【００１９】本発明の他の目的は、データ多重化を行うマスタシステムおよびリモートシステムにおいて、マスタシステム側の障害等に際してリモートシステムに運用を切り替える上で運用可能な整合性の保証されたデータを提供することにある。

【００２０】

【課題を解決するための手段】本発明は、第１のシステムに設けられた第１の記憶装置に格納されたデータを、少なくとも一つの第２のシステムに設けられた第２の記憶装置に複写することでデータの多重化を行うデータ多重化方法において、任意の要因にて第１のシステム側に蓄積された未複写のデータを第２のシステム側に転送する回復複写処理（「回復コピー」）の実行に際して、第２のシステム側では、第１のシステムから到来するデータに対応した旧データを任意の方法にて保存しておき、回復複写処理が失敗した時には、保存されている旧データを用いて第１のシステムと第２のシステムとの間におけるデータの整合性を維持するものである。

【００２１】すなわち、より具体的には、一例として、「同期方式」のデータ多重化にてマスタシステム（第１のシステム）とリモートシステム（第２のシステム）との間におけるデータの「ミラー状態」を維持する場合に、リモートシステム側において、マスタシステム側と「ミラー状態」を維持していた最近、即ち「サスペンド状態」に陥った直前のデータを、「回復コピー」実行の間、別の記憶領域に一時的に退避して保存するものである。

【００２２】なお、旧データの保存方法としては、「回

復コピー」実行の間にリモートシステムに到来するバックアップ対象データを、「回復コピー」が完了するまでの間は対応する旧データの上書きせずに、別の記憶領域に一時的にストアし、リモートシステム側の第２の記憶装置上には「サスペンド状態」に陥った直前の整合性を維持した状態の旧データを残しておく方法でもよい。

【００２３】また、「回復コピー」中の障害の発生に備えて、マスタシステム側では、リモートシステムが「サスペンド状態」に陥ってから「回復コピー」実行に至るまでの各「差分データ」を、例えば時系列キューを構築して管理することにより発生順序に従って保持する。そして、最も過去にキューに組み入れた差分データから、「回復コピー」を始めることによって、リモートシステム側は、特に差分データの整合性を意識することなしに、「回復コピー」に障害が発生した場合には、当該障害直前の旧データを採用することで、バックアップデータの整合性を保つことができる。

【００２４】また、「回復コピー」中におけるマスタシステムおよびマスタシステムとリモートシステムとの間におけるデータ転送経路等の障害の有無をリモートシステム側から的確に判別すべく、「回復コピー」の完了までの許容最大時間を設定する機能を設け、予め、設定された許容最大時間を超過してもマスタシステム側から「回復コピー」の完了報告がない場合には、リモートシステム側では、マスタシステムが重大障害に陥り、運用不可の状態になっていると判断して、保存されていた旧データのレベルで整合性をとる等の整合性復旧のための処理を自動的に開始する構成とすることもできる。

【００２５】

【発明の実施の形態】以下、本発明の実施の形態を図面を参照しながら詳細に説明する。

【００２６】図１は、本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムの構成の一例を示す概念図である。

【００２７】本実施の形態の情報処理システムは、マスタシステム１００（Ｍ－ＳＹＳ）と、このマスタシステム１００が取扱うデータのバックアップシステムとして機能するリモートシステム１１０（Ｒ－ＳＹＳ）からなる。

【００２８】マスタシステム１００は、ホストＣＰＵ１０１、ディスク制御装置１０２およびディスク駆動装置１０３の組み合わせからなる情報記憶システム一系列を構成しており、同様にリモートシステム１１０においても、ホストＣＰＵ１１１、ディスク制御装置１１２およびディスク駆動装置１１３の組み合わせからなる情報記憶システム一系列を構成している。

【００２９】ディスク制御装置１０２とディスク制御装置１１２の間は、各々に設けられたコネクタ１０６，コネクタ１１６を介して、例えば専用回線や公衆回線等の情報通信網等からなるデータ転送路１２０にて接続されている。

【００３０】ディスク駆動装置１０３、ディスク駆動装置１１３としては、一例として、例えば磁気ディスク、光ディスク、光磁気ディスク、デジタルビデオディスク（ＤＶＤ）等、任意の記憶媒体を用いた記憶装置を用いることができる。

【００３１】ホストＣＰＵ１０１は、ディスク駆動装置１０３のディスク媒体へのデータ書込み命令の発行をはじめ、ディスク制御装置１０２を制御する為の指示を発行し、ディスク制御装置１０２は、ホストＣＰＵ１０１からの指示を解析し、その結果に基づいて制御を行なう。例えば、ホストＣＰＵ１０１よりデータの書込み指示があり、そのデータが送られてくると、ディスク制御装置１０２はその指示に従い、目的のディスク駆動装置１０３にデータを書込む。更に、ディスク制御装置１０２は、リモートシステム１１０の側に向けて、データのバックアップに関するリモートシステム１１０への制御指示、およびバックアップデータを転送する機能を有する。また、リモートシステム１１０側のディスク制御装置１１２は、マスタシステム１００からの指示やバックアップデータの受付けを実施し、書込みデータをディスク駆動装置１１３のディスク媒体上に反映する機能を有する。

【００３２】また、マスタシステム１００のディスク制御装置１０２およびディスク駆動装置１０３、そして、リモートシステム１１０のディスク制御装置１１２およびディスク駆動装置１１３は、各々、キャッシュメモリ１０４，キャッシュメモリ１１４、更には、ディスク制御装置１０２、ディスク制御装置１１２に対する様々な情報や機能（例えばバックアップコピーの為の各種条件やシステム内の構成情報等の設定や指示機能）を提供するサービスプロセッサ１０５（ＳＶＰ），サービスプロセッサ１１５（ＳＶＰ）を持つ。

【００３３】また、本実施の形態の場合、一例として、ディスク駆動装置１０３およびディスク駆動装置１１３は、ＲＡＩＤ５のアーキテクチャを有する。

【００３４】尚、本実施の形態においては、一例として上述の通り構成されるが、本発明においては、記憶装置としては、ディスク駆動装置やＲＡＩＤ５を採用することに限らず、一般の記憶装置（ＲＡＩＤアーキテクチャを有さないディスク装置、半導体記憶装置、磁気テープ装置など）や他のＲＡＩＤレベル（ＲＡＩＤ０，ＲＡＩＤ１，．．．）構成に対しても適用可能である。また、ホストＣＰＵ１０１，１１１、およびディスク制御装置１０２，１１２、更にＳＶＰ１０５，１１５における上述の機能については、各々マスタシステム１００および、リモートシステム１１０内に設けられているならば代用可能である。例えば、ディスク制御装置１０２、１１２が自らデータの書き込み指示を発行して、ディスク媒体上への反映を行なったり、バックアップコピーに関

する条件設定をＳＶＰを介してではなく、ホストＣＰＵ１０１やディスク制御装置１０２自身で設定するケースなどが考えられる。

【００３５】図２は、本実施の形態のディスク駆動装置１０３およびディスク駆動装置１１３の各々におけるＲＡＩＤ５のアーキテクチャの一例を示す概念図である。

【００３６】ＲＡＩＤ５のアーキテクチャを有するディスク制御装置／駆動装置とは、多数の小型ディスク２００，２０１，．．．を記憶媒体として用い、大型ディスク装置をエミュレイトしたしたディスク装置であり、何本かの論理トラック２１０，２１１，．．．に対して、データ修復コードであるパリティコードを格納する１本の論理トラック（パリティトラック２２０）を用意しており、ある１台の小型ディスクが故障してもパリティトラック２２０のパリティデータを元に故障ディスク内の全データの修復を自発的に行い、故障したディスクの代替用の小型ディスク（スペアディスク２３０）にその回復データを反映することにより、オンラインでのシステム運用継続を可能にするなどの特徴を持っている。

【００３７】以上に述べた構成の本実施の形態の情報処理システムにおいて、例えば、整合性を維持する単位として、論理デバイスを選んだ場合のバックアップコピーの手順の一例を説明する。図３は、その手順の一例を図式化して例示した概念図であり、図４はそのフローチャートである。

【００３８】マスタシステム１００において、ホストＣＰＵ１０１、或いはＳＶＰ１０５などより、論理デバイス単位でマスタシステム１００とリモートシステム１１０の間でペア（マスタシステム１００側の１論理デバイス（Ｍ－ＤＥＶ）とリモートシステム１１０側の１論理デバイス（Ｒ－ＤＥＶ）の組であり、Ｒ－ＤＥＶは、Ｍ－ＤＥＶに対するデータのコピー先にあたり、Ｍ－ＤＥＶのデータは、指定されたＲ－ＤＥＶ内に格納される。）３１０，３１１，．．．，３１２を形成するよう指示が発生すると（ステップ３５０）、マスタシステム１００は（ディスク制御装置１０２等の記憶制御装置などを介して）、リモートシステム１１０に向けて、その旨を指示する命令を発行する（ステップ３５１）。その正常終了報告をリモートシステム１１０側より受けると（ステップ３５２）、マスタシステム１００はバックアップコピーを開始し、Ｍ－ＤＥＶ＃＝Ｍｉに入った更新データは、そのペアにあたるＲ－ＤＥＶ＃＝Ｒｉに反映を行い、「ミラー状態」を維持する（ステップ３５３）。なお、この「ミラー状態」とは、あくまでも、真のミラー状態である必要はなく、データの多重化が維持されていればよく、物理的に全く同一のトラックに反映しなくてもかまわない。

【００３９】ところが、ケーブル等のデータ転送路１２０が切断されたり、マスタシステム１００側のディスク駆動装置１０３の媒体障害などでペア間のデータコピーが一時的に不可能となった、即ち、ペアが「サスペンド状態」に陥った場合、ケーブルの交換、ディスク駆動装置１０３のオンライン交換（データはパリティデータより再生成可能）などにより、障害要因が取り除かれ、ペア間のデータコピーが再開可能になるまでの間、ホストＣＰＵ１０１より発行される更新Ｉ／Ｏに対し、ディスク制御装置１０２は更新のあった論理デバイス番号（Ｌ－ＤＥＶ＃）および論理トラック番号（Ｌ－ＴＲＫ＃）を後述のような差分ビットマップ１０７を利用して記憶する。

【００４０】図７は、この差分ビットマップ１０７の構成の一例を示す概念図である。ここでは、論理デバイス番号を行（論理デバイス＃＝０を１行目、．．．論理デバイス＃＝ｉをｉ＋１行目）、論理トラック番号を列（論理トラック＝０を１列目、．．．論理トラック＃＝ｊをｊ＋１列目）に配列したものとなっている。例えば、マスタシステム１００側にて、論理デバイス＃＝ｉの論理トラック＃＝０に差分（トラック）データが発生すると、（ｉ＋１）行、１列目のビット１０７ａを立てることにより差分が当該トラックに存在することを表わしている。

【００４１】やがて、ペア間のデータコピーが再開可能となった後、ホストＣＰＵ１０１からの指示、或はＳＶＰ１０５からの指示によって、「回復コピー」を始める。

【００４２】最初に、この「回復コピー」において、リモートシステム１１０側では、マスタシステム１００側から転送されて来るデータに対応したディスク駆動装置１１３内の旧データを、上書きされる前に、他の領域に退避させて保存する場合について説明する。図５および図６は、この場合の「回復コピー」の処理の流れの一例を示したフローチャートであり、図５はマスタシステム側の動作を、図６はリモートシステム側の動作を、それぞれ示している。

【００４３】本実施の形態においては、旧データの退避エリアとして、リモートシステム１１０配下のディスク制御装置１１２に搭載されているキャッシュメモリ１１４および、ディスク駆動装置１１３内のパリティトラック２２０（パリティエリア）或は、スペアディスク２３０の記憶領域を利用している。特にパリティトラック２２０を退避エリアとして使用することはリモートシステム１１０側のディスク駆動装置１１３が、「回復コピー」を実施したことを契機にＲＡＩＤ５の機能レベルを失うことを意味するが、それよりも、「回復コピー」中にマスタシステム１００が運用不可に陥り、更にリモートシステム１１０側のディスクデータも保証出来ないという技術的課題を解決する方がはるかに大切である。もちろん、パリティトラック２２０やスペアディスク２３０の使用可否等の条件をユーザ側にて選択できる機能は用意する。但し、一般的には、「回復コピー」が終了し

て、各ペアが「ミラー状態」を回復した瞬間に至れば、マスタシステム１００のパリティデータを考慮に入れれば、リモートシステム１１０側のディスク駆動装置１１３のＲＡＩＤ５レベルは確保できる。

【００４４】さて、マスタシステム１００配下のディスク制御装置１０２は回復コピー指示のイベントを受けると（ステップ４０１）、まず、ＳＶＰ１０５から回復コピー完了許容最大時間Ｔを入手し（ステップ４０２）、さらに、リモートシステム１１０側の記憶制御装置（ディスク制御装置１１２）のキャッシュメモリ容量、パリティトラック本数およびスペアディスク数などの構成情報をリモートシステム１１０より入手する（ステップ４０３）。

【００４５】さらに、ユーザが設定した条件（ＲＡＩＤ５レベルの維持、スペアディスク２３０の使用可否、キャッシュメモリ１１４の最大占有許容量（しきい値）などがあり、リモートシステム１１０側よりこれらの条件を設定する場合、マスタシステム１００は「回復コピー」を開始する前に、予め入手しておく。また、マスタシステム１００側よりこれらの情報を設定する場合は、リモートシステム１１０側に、これらの情報を与えておき、これらの条件を双方のシステムが同一認識している中で「回復コピー」を実施する）を元に、旧データの退避エリアの容量を算出し、同時に「回復コピー」を実施することが可能なペア数の最大値ｎを決定する（ステップ４０４）。

【００４６】この最大値ｎの決定の理由は、「回復コピー」処理の効率の面から、確保できる旧データの退避エリアの容量に対応した推奨値を提供する目的で求めるものであり、単に１論理デバイス単位毎に「回復コピー」を実施しても構わない。

【００４７】例えば、リモートシステム１１０側のディスク駆動装置１１３の構成例として、６４論理デバイスを２４台の物理ディスク（小型ディスク）でエミュレイトし、小型ディスク４台分のパリティトラック（４パリティグループ）および２台のスペアディスク２３０からなり、キャッシュメモリ１１４は考慮にいれない場合、ｎとして設定できる最大値は（６４／２４）×（４＋２）の整数部分をとったもの、即ち、ｎ＝１６であり、「回復コピー」を実施する上で、旧データの退避エリアがネックとなることはないといえる。また、キャッシュメモリ１１４を充分に搭載したシステム環境を整えることが可能ならば、キャッシュメモリ１１４のみを退避エリアとして使用することも可能である。

【００４８】次に、マスタシステム１００側で「回復コピー」の為の準備が整った段階で、リモートシステム１１０側に「回復コピー」開始を指示し（ステップ４０５）、「回復コピー」にあたって使用する情報（回復コピー完了許容最大時間：Ｔなど）を送り、指示に対する応答を待つ（ステップ４０６）。

【００４９】一方、リモートシステム１１０側では、「回復コピー」の指示を受けると（ステップ５０１）、「回復コピー」開始を認識し、監視タイマ（回復コピー完了許容最大時間Ｔに対する監視）をスタートさせるなどの準備を行い（ステップ５０２）、“回復コピー中”のフラグをＯＮにし（ステップ５０３）、マスタシステム１００に準備が整ったことを表わす応答を返し（ステップ５０４）、差分データの到着を待つ（ステップ５０５）。

【００５０】さて、リモートシステム１１０より「回復コピー」開始指示に対する応答を受けたマスタシステム１００側は、差分ビットマップ１０７（図７）を参照することにより、差分データがあるトラック（差分トラック）を検索し（ステップ４０７）、見つかると、そのトラックデータをリモートシステム１１０側に送るべく、書き込み命令を作成・発行し（ステップ４０８）、リモートシステム１１０からの書き込み終了報告を待つ（ステップ４０９）。

【００５１】その書き込み命令受信を検知すると、リモートシステム１１０側は、バックアップドライブ（ディスク駆動装置１１３）への格納アドレスを求め（ステップ５０６）、キャッシュメモリ１１４上に差分データを格納し（ステップ５０７）、書き込み終了報告を行う（ステップ５０８）。

【００５２】その後、この差分データをディスク駆動装置１１３の目的のトラックに上書きする前に、当該トラック上の旧データを退避するために、空きエリア（トラック）を探したのち（ステップ５０９）、一旦トラック上の旧データをキャッシュ上に退避し、旧データが元格納されていたデバイス番号やトラック番号および退避先などの情報を作成して、記憶しておく（ステップ５１０）。

【００５３】尚、退避データが使用可能キャッシュ許容量を越えない限りは（ステップ５１１）、キャッシュメモリ１１４のみを使用して退避を続け、そのしきい値を越えた場合に限って、ディスク駆動装置１１３上の退避エリアに反映する（ステップ５１２）。

【００５４】以上が、１本分の差分トラックに対する「回復コピー」の処理になるが、このあとも引き続き、マスタシステム１００側は、差分トラックを検索して（ステップ４１０）、リモートシステム１１０側に書き込み命令を送り、あるペアの全差分データを反映し終わると、そのペアｉに対する回復到達報告をリモートシステム１１０側に行い（ステップ４１１）、「回復コピー」待ちになっていたペアを追加して（ステップ４１２）、「回復コピー」指示のあった全てのペアに対する差分データをリモートシステム１１０側に送り続ける。そして、全ての差分データがリモートシステム１１０側に反映されたところで「回復コピー」完了報告をリモートシステム１１０に送り（ステップ４１３）、マスタシ

ステム１００側の「回復コピー」に対する処理は完了する。

【００５５】さて、リモートシステム１１０側は、「回復コピー」中は、マスタシステム１００側からの書き込み命令受信、マスタシステム１００側からの「回復コピー」完了報告および、「回復コピー」実行時間タイムオーバのいずれかを検知すると、それに対応した処理へ移行するようになっており（ステップ５０５、ステップ５１４、ステップ５１５で形成されるループ）、「回復コピー」完了報告を受けると、リモートシステム１１０側は、「回復コピー」完了を認識して“回復コピー中”のフラグをＯＦＦし（ステップ５２０）、「回復コピー」処理を完了させる。

【００５６】一方、マスタシステム１００側からの「回復コピー」完了報告がいつまでたってもなく、「回復コピー」実行タイムオーバに達した場合においては、回復到達報告（図５のステップ４１１）があったペア以外の、整合性を保証できないリモートシステム１１０側の論理デバイスに対する退避データの有無をチェックし（ステップ５１６）、退避データがキャッシュメモリ１１４上に残っているか否かを判別し（ステップ５１７）、残っていればキャッシュメモリ１１４より（ステップ５１８）、そうでない退避データは、パリティトラック２２０やスペアディスク２３０からキャッシュメモリ１１４上に読み込みを行なって、バックアップを行なっている目的のトラックに旧データを反映し（ステップ５１９）、整合性復旧の為の処理を実施する。

【００５７】次に、旧データの退避方法、および手順の一例について詳細に説明を行なう。図８は、その手順の一例を図式化して例示した概念図であり、図９は、そのフローチャートである。

【００５８】まず、リモートシステム１１０は、マスタシステム１００から差分データが転送され、そのデータをディスク制御装置１１２のキャッシュメモリ１１４内に格納すると、差分データ格納終了報告をマスタシステム１００に対し行なう（ステップ７５１、ステップ７５２）。その後、差分データによって上書きされる旧データを退避する為に、システム管理用メモリ１１７上の退避エリア管理ビットマップ１１８をアクセスし、空きトラックを検索、退避先を決定する（ステップ７５３）。

【００５９】なお、システム管理用メモリ１１７の割り当て領域には特に制限はなく、リモートシステム１１０内の任意のメモリ（例えばキャッシュメモリ１１４など）上に任意に定義することができる。

【００６０】図１０は、退避エリア管理ビットマップ１１８の具体例を例示した概念図である。パリティトラック群、あるいは、スペアディスク１台分を行に、およびトラック番号を列に配列し、各退避先トラック１本１本の使用状況をビットマップ８０２で表現したものである。場合によっては、パリティトラック２２０、スペアディスク２３０の使用可否フラグ８０１（ビットマップの１列目に配置）を持つこともある。図１０の例では、スペアディスク＃０に対する使用可否フラグ８０３が立っている為、スペアディスク＃０のトラックが退避先として許可されており、更に、その論理トラック番号０、ｎに対するスペアディスク＃０上のエリアは“空き”であることを表わす。

【００６１】退避エリアを確保すると、差分データによって上書きされる旧データを退避データとしてキャッシュメモリ１１４上に読み込み（ステップ７５４）、併せて、整合性回復実施時に必要な退避データに関する諸情報を作成し、１つのノードとして、システム管理用メモリ１１７の一部に設けられた退避データ情報キュー１１９にエントリする（ステップ７５５）。

【００６２】図１１は、この退避データ情報キュー１１９の構造と各ノードに関する諸情報の一例を表わす概念図である。この例では、差分データの反映先である論理デバイスの番号毎に、１つのキュー９０１、９０２、．．．，９０３を設け、１論理トラック分を単位として１つのノードを割り当てている。尚、本実施の形態では一例として退避データ情報キュー１１９のようなキュー構造を採用しているが、本発明においては、単純なテーブル構造等、任意のデータ構造を採用することも可能である。

【００６３】各ノードには、整合性回復を行なう際に必要な情報、具体的には差分データの反映先にあたる物理デバイス番号９５０やその格納アドレス（物理トラック番号（セクタ番号）９５１）、旧データの退避先にあたる（スペア）ディスク番号９５２、および物理トラック番号（セクタ番号）９５３等の格納アドレス、および旧データのキャッシュメモリ上格納アドレス９５４などを含んでいる。

【００６４】そして、最後に、差分データを記憶媒体（ディスク駆動装置１１３）上に、退避旧データをパリティトラック２２０やスペアディスク２３０上の退避エリアに反映し、１差分データに対する、旧データの退避は完了する（ステップ７５６、ステップ７５７）。なお、このステップ７５６、ステップ７５７の操作は、非同期に実施されても構わない。

【００６５】次に、リモートシステム１１０側における「回復コピー」実行時の旧データの保存方法の他の例を説明する。この例では、「回復コピー」実行時に旧データの退避は行わず、マスタシステム１００から到来する差分データを本来のディスク駆動装置１１３の書込領域以外の領域に一時的に保持することで、旧データの保存を行う。この例について、図６を参照して、前述の実施の形態との違いを挙げながら説明を行う。

【００６６】違いは３つあり、１つ目は、退避エリアを確保すべく、退避エリア管理ビットマップ１１８を使って、空きエリア（トラック）を探した後（ステップ５０

９に対応）、その次のステップ５１０は行わずに、退避データ情報キュー１１９に差分データに対する情報を作成しエントリを行う（ステップ５１１に対応）。本実施の形態においても、キャッシュメモリ占有許容量（しきい値）を越えない限りは、基本的に退避データ（差分データ）をキャッシュメモリ１１４上に置いておく。

【００６７】２つ目は、「回復コピー」完了報告後の処理であり、全差分データを目的のディスク駆動装置１１３に反映する必要がある。キャッシュメモリ１１４上になくパリティトラック２２０やスペアディスク２３０上に退避してある差分データは、キャッシュメモリ１１４の空き容量に応じてキャッシュメモリ１１４上に読み込みを行い、順次、目的のバックアップを行うべきディスク駆動装置１１３に差分データの反映を行う。

【００６８】また、３つ目は、「回復コピー」実行タイムオーバになったケース、即ち、整合性維持のための処理の部分であり、回復到達報告を受けたペアに対しては、差分データの反映を実施するが、回復到達報告を受けていないペアに対しては差分データは目的のバックアップを行うべきディスク駆動装置１１３への反映を行わずに破棄する。

【００６９】次に、マスタシステム１００の側において「サスペンド状態」の時に自システム内で発生した更新データ（差分データ）を時系列に蓄積する場合に用いられる時系列キューの構築方法の一例について説明する。

【００７０】図１２は、その時系列キュー１０８の構造を表わしており、１００１、１００２、．．．は時系列キュー１０８の各ノードにあたる。この時系列キュー１０８は、マスタシステム１００内にあるメモリ（例えばディスク制御装置１０２のキャッシュメモリ１０４や図示しないシステム管理用メモリなど）に定義・構築し、少なくとも整合性の維持を図る単位（例えば論理デバイスなど）毎に設ける。各ノードは、差分データの発生順にキュー先頭から並んでおり、例えば、論理トラック単位毎に１個のノードを割り当てる。

【００７１】時系列キュー１０８を構成する個々のノード１００１、ノード１００２、．．．は、直前に入った差分データに対する、今まで最後尾に接続されていたノードの先頭をさす前エントリノードポインタ１１０１と、次に発生した差分データに対するノードの先頭をさす次エントリノードポインタ１１０６とを有する双方向キューの形式をとっている。なお、前エントリノードポインタ１１０１を省略した一方向キューの形式を採ってもよいことは言うまでもない。

【００７２】更に、個々のノード１００１、ノード１００２、．．．は、本ノード制御用情報１１０２と、当該ノードに割り当てられた差分データの論理トラック番号１１０３と、差分データの格納先を表わす物理媒体上格納アドレス１１０４と、キャッシュメモリ上アドレス１１０５とを含んでいる。

【００７３】本ノード制御用情報１１０２には、例えば当該ノードに対する差分データが、論理トラック全体をフォーマットするための更新（フォーマットライト）データなのか、部分的な更新（パーシャルライト）データなのかを表わす更新タイプ情報１１０２ａや、当該ノードに対する差分データはリモートシステム１１０側に転送すべきかどうかを表わす本ノード有効フラグ１１０２ｂなどを含んでいる。

【００７４】次に、上述のような構成の時系列キュー１０８へのエントリ手順の一例を、図１３フローチャートを参照して説明する。ここでは、ある論理デバイスの特定の論理トラックへの差分データに対するノードのキューへのエントリに対する例を挙げており、その更新データがマスタシステム１００側で発生すると（ステップ１２０１）、そのデータの格納アドレス（キャッシュメモリ上アドレス、物理記憶媒体上（物理デバイス）アドレス）を求め（ステップ１２０２）、更新データをキャッシュメモリ１０４などに格納する（ステップ１２０３）。その後、本更新データに対する情報を組み合わせてノード情報を作成して（ステップ１２０４）、目的の時系列キュー１０８に接続する（ステップ１２０５）。更に、本論理トラックに対するフォーマット形式の差分データが既にあったか否かを調べ（ステップ１２０６）、あった場合には、以前に発生したフォーマットデータにあたる差分データはリモートシステム１１０側に転送する必要はないため、以前にエントリしたフォーマット形式の差分データに対するノードを検索し、そのノードに対する本ノード有効フラグ１１０２ｂをＯＦＦにするか、あるいは時系列キュー１０８からはずすなどの手順を実施し（ステップ１２０７）、差分ビットマップへの反映は行わない。また、前記ステップ１２０６で、なかったと判定された場合には、差分ビットマップ１０７への反映を行う（ステップ１２０８）。

【００７５】このような時系列キュー１０８を使用して（必要に応じて差分ビットマップを併用して）、回復コピー開始に伴い、その先頭位置にあるノードに対する差分データ、すなわち最も過去に発生した差分データから時系列に転送を開始することにより、整合性を維持すべき単位（論理デバイス単位など）での整合性を常に維持することが可能になる。

【００７６】以上説明したように、本実施の形態のデータ多重化方法によれば、例えば「同期型」のデータ多重化を行う情報処理システム等において、なんらかの障害で「サスペンド状態」に陥った後の「回復コピー」実行中にマスタシステム１００が重大災害によって運用不可な状態に陥ったとしても、リモートシステム１１０側のバックアップデータに対する整合性の保証を保持することができる為、リモートシステム１１０への運用移行を的確に行うことができる。

【００７７】以上本発明者によってなされた発明を実施

の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【００７８】たとえば、上述の実施の形態の説明では、一例としてマスタシステムおよびリモートシステムの二つのシステム間でデータを二重に保持させる例を説明したが、３以上のシステムにてデータ多重化を行う場合にも、本発明は適用できる。

【００７９】

【発明の効果】本発明のデータ多重化方法によれば、複数のシステム間でのデータ多重化を維持するための差分データの回復複写における障害耐性を向上させることができる、という効果が得られる。

【００８０】また、本発明のデータ多重化方法によれば、マスタシステムおよびリモートシステムが、ある整合性の維持を要求する情報単位において、特に「サスペンド状態」に陥る瞬間まで「ミラー状態」を維持していた場合の、「回復コピー」におけるリモートシステム側の多重化データの整合性を保証することができる、という効果が得られる。

【００８１】また、本発明のデータ多重化方法によれば、マスタシステム側の障害等に際してリモートシステムに運用を切り替える上で運用可能な整合性の保証されたデータを提供することができる、という効果が得られる。

【図面の簡単な説明】

【図１】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムの構成の一例を示す概念図である。

【図２】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおけるＲＡＩＤ５のアーキテクチャの一例を示す概念図である。

【図３】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムでのバックアップコピーの手順の一例を図式化して例示した概念図である。

【図４】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムでのバックアップコピーの手順の一例を示すフローチャートである。

【図５】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおけるマスタシステムの作用の一例を示すフローチャートである。

【図６】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおけるリモートシステムの作用の一例を示すフローチャートである。

【図７】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにて使用される差分ビットマップの一例を示す概念図である。

【図８】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおける回復コピー手順の一例を図式化して例示した概念図である。

【図９】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおける回復コピー手順の一例を示すフローチャートである。

【図１０】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおける退避エリア管理ビットマップの一例を例示した概念図である。

【図１１】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおける退避データ情報キューの構造の一例を示す概念図である。

【図１２】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおける時系列キューの構造の一例を示す概念図である。

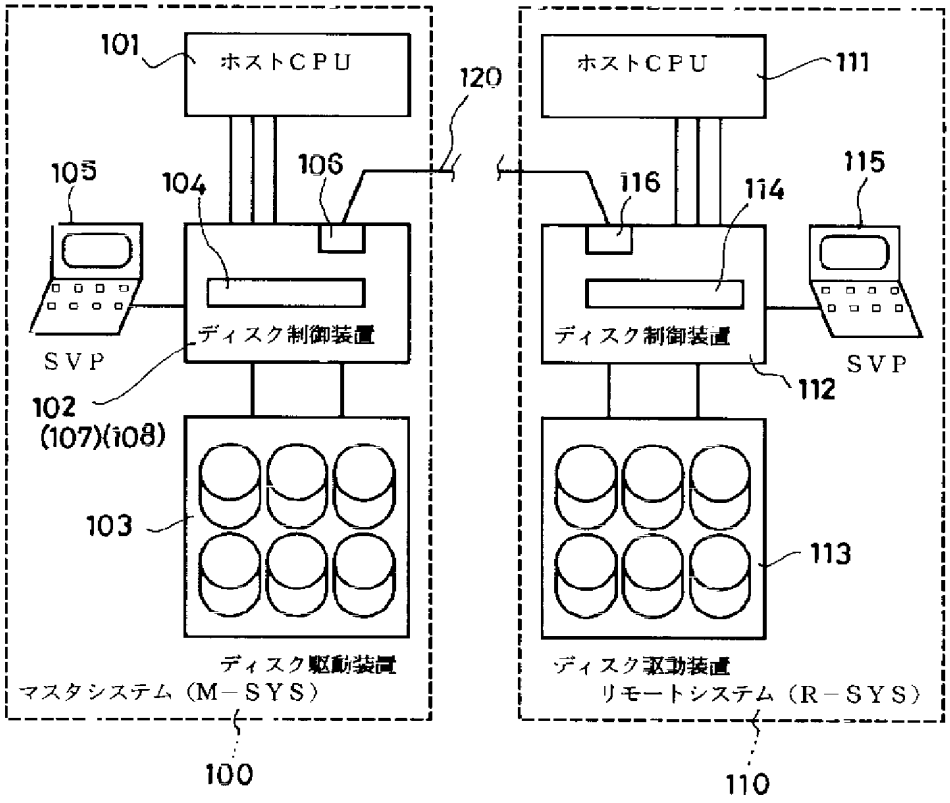
【図１３】本発明の一実施の形態であるデータ多重化方法が実施される情報処理システムにおける時系列キューのエントリ手順の一例を示すフローチャートである。

【符号の説明】

１００…マスタシステム（第１のシステム）、１０１…ホストＣＰＵ、１０２…ディスク制御装置、１０３…ディスク駆動装置（第１の記憶装置）、１０４…キャッシュメモリ、１０５…サービスプロセッサ、１０６…コネクタ、１０７…差分ビットマップ、１０８…時系列キュー、１１０…リモートシステム（第２のシステム）、１１１…ホストＣＰＵ、１１２…ディスク制御装置、１１３…ディスク駆動装置（第２の記憶装置）、１１４…キャッシュメモリ、１１５…サービスプロセッサ、１１６…コネクタ、１１７…システム管理用メモリ、１１８…退避エリア管理ビットマップ、１１９…退避データ情報キュー、１２０…データ転送路、２００…小型ディスク、２１０…論理トラック、２２０…パリティトラック、２３０…スペアディスク、８０１…使用可否フラグ、８０２…ビットマップ、９５０…物理デバイス番号、９５１…物理トラック番号（セクタ番号）、９５２…ディスク番号、９５３…物理トラック番号（セクタ番号）、９５４…キャッシュメモリ上格納アドレス、１１０１…前エントリノードポインタ、１１０２…本ノード制御用情報、１１０２ａ…更新タイプ情報、１１０２ｂ…本ノード有効フラグ、１１０３…論理トラック番号、１１０４…物理媒体上格納アドレス、１１０５…キャッシュメモリ上アドレス、１１０６…次エントリノードポインタ、Ｔ…回復コピー完了許容最大時間。

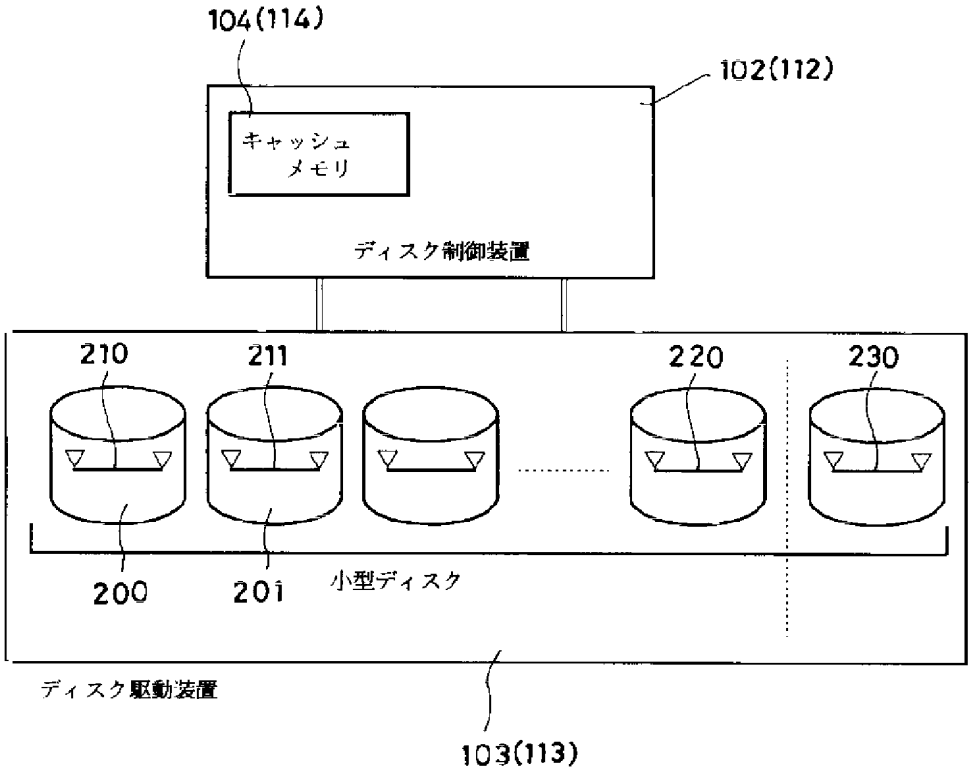
【図１】

図１


101
ホストＣＰＵ
120
ホストＣＰＵ
111
105
104
106
116
114
115
ディスク制御装置
ディスク制御装置
ＳＶＰ
ＳＶＰ
102
(107)(108)
112
103
113
ディスク駆動装置
マスタシステム（Ｍ－ＳＹＳ）
ディスク駆動装置
リモートシステム（Ｒ－ＳＹＳ）
100
110

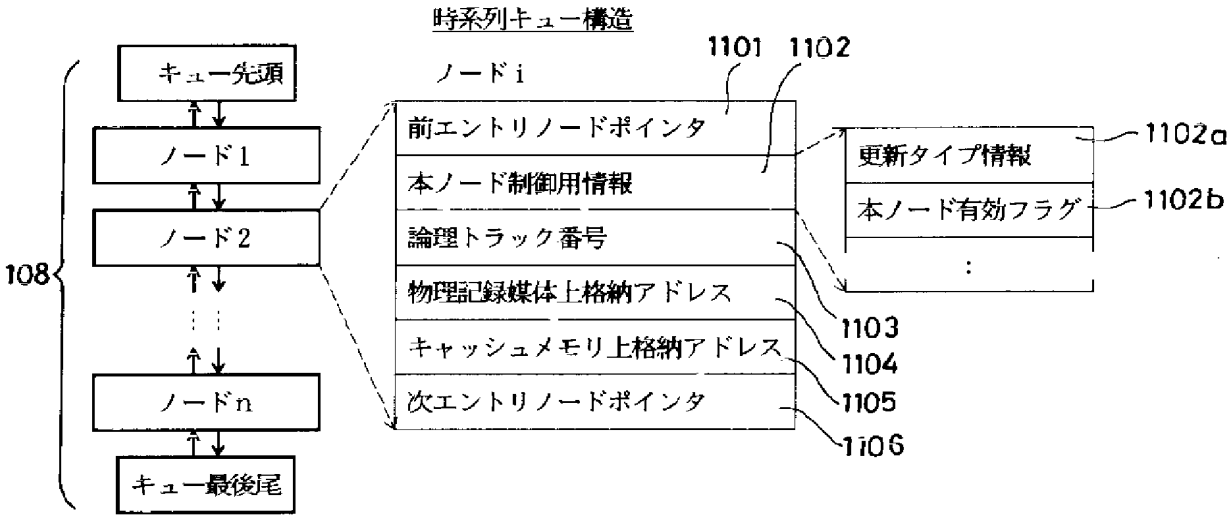
【図２】

図２


104(114)
102(112)
キャッシュ
メモリ
ディスク制御装置
210
211
220
230
200
201
小型ディスク
ディスク駆動装置
103(113)

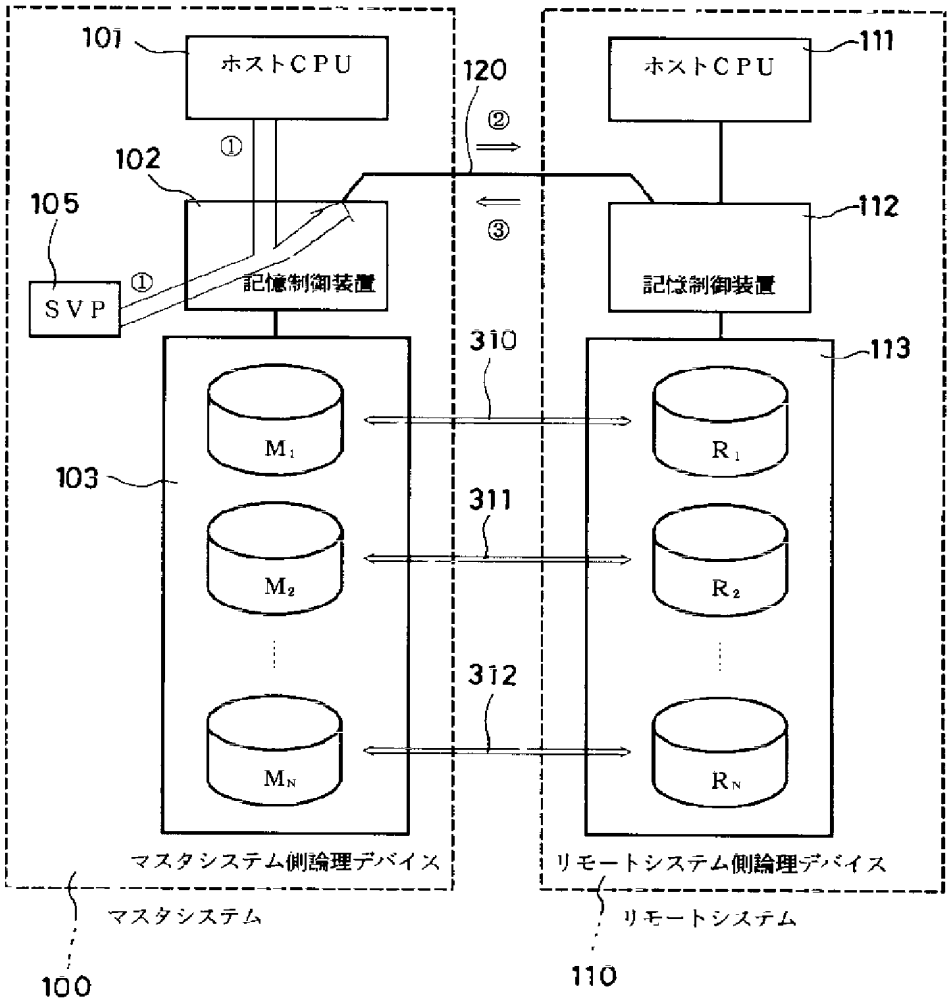
【図１２】

図１２


時系列キュー構造
キュー先頭
ノード1
ノード2
ノードn
キュー最後尾
108
ノードi
1101
1102
前エントリノードポインタ
本ノード制御用情報
論理トラック番号
物理記録媒体上格納アドレス
キャッシュメモリ上格納アドレス
次エントリノードポインタ
1103
1104
1105
1106
更新タイプ情報
本ノード有効フラグ
1102a
1102b

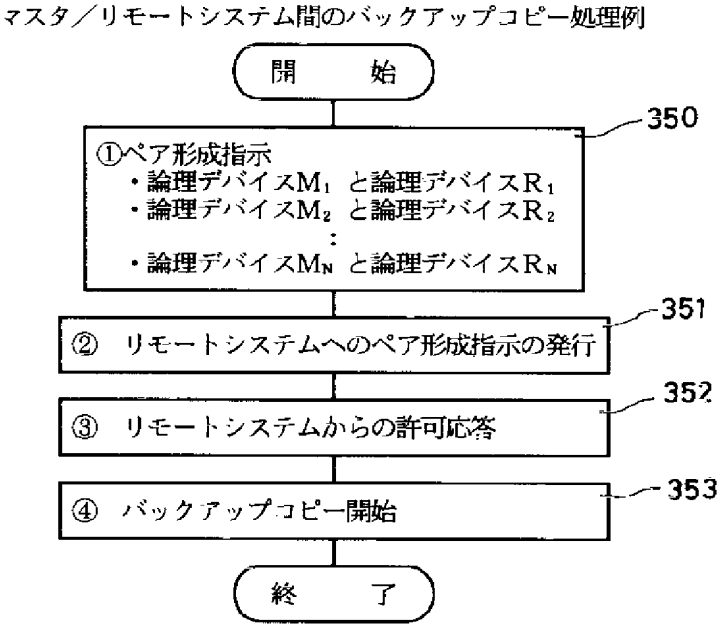
【図３】

図３


101
ホストＣＰＵ
120
②
111
ホストＣＰＵ
①
102
105
③
112
①
ＳＶＰ
記憶制御装置
記憶制御装置
310
113
M_1
R_1
103
311
M_2
R_2
312
M_N
R_N
マスタシステム側論理デバイス
リモートシステム側論理デバイス
マスタシステム
リモートシステム
100
110

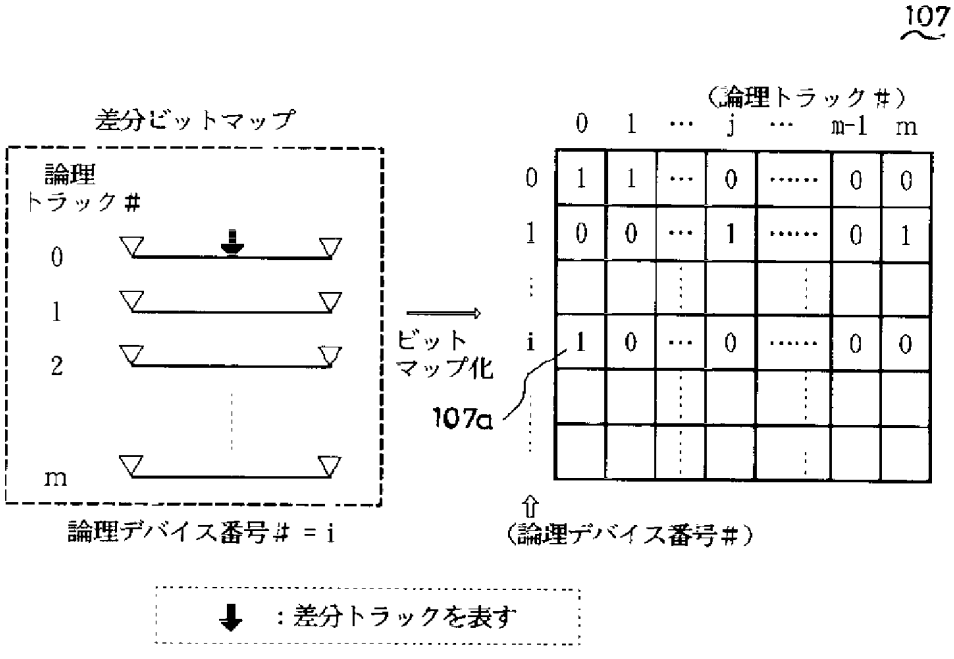
【図４】

図４


マスタ／リモートシステム間のバックアップコピー処理例
開　始
350
①ペア形成指示
・論理デバイスM1 と論理デバイスR1
・論理デバイスM2 と論理デバイスR2
：
・論理デバイスMN と論理デバイスRN
351
② リモートシステムへのペア形成指示の発行
352
③ リモートシステムからの許可応答
353
④ バックアップコピー開始
終　了

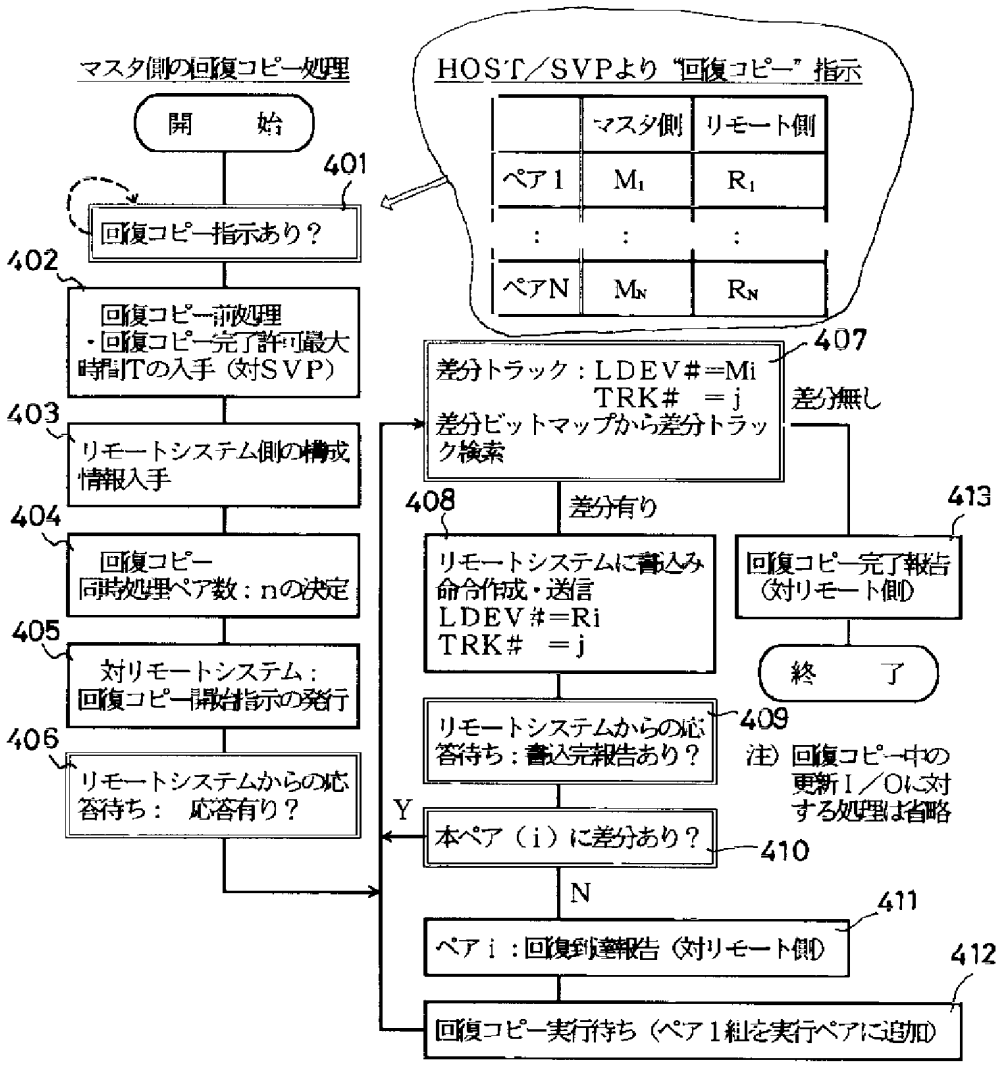
【図７】

図７


107
差分ビットマップ
論理
トラック＃
0
1
2
m
論理デバイス番号＃ = i
ビット
マップ化
107a
（論理トラック＃）
0 1 … j … m-1 m
0 1 1 … 0 …… 0 0
1 0 0 … 1 …… 0 1
i 1 0 … 0 …… 0 0
⇧
（論理デバイス番号＃）
：差分トラックを表す

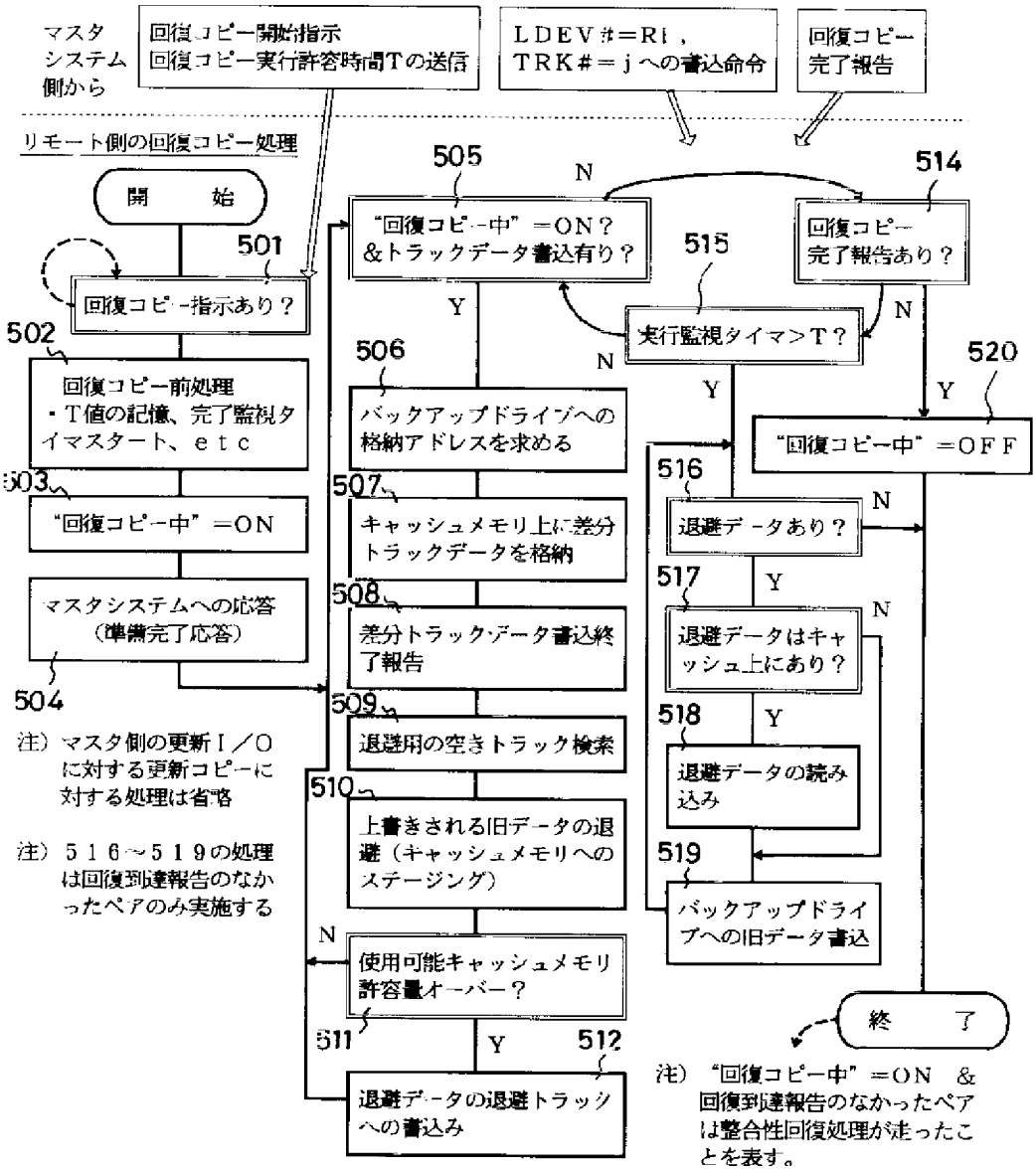
【図５】

図5


マスタ側の回復コピー処理
開　始
401
回復コピー指示あり？
402
回復コピー前処理
・回復コピー完了許可最大時間Tの入手（対ＳＶＰ）
403
リモートシステム側の構成情報入手
404
回復コピー
同時処理ペア数：nの決定
405
対リモートシステム：
回復コピー開始指示の発行
406
リモートシステムからの応答待ち：　応答有り？
HOST／SVPより“回復コピー”指示
マスタ側
リモート側
ペア1
M1
R1
ペアN
MN
RN
407
差分トラック：ＬＤＥＶ＃＝Mi
ＴＲＫ＃　＝ｊ
差分ビットマップから差分トラック検索
差分無し
408
差分有り
リモートシステムに書込み命令作成・送信
ＬＤＥＶ＃＝Ri
ＴＲＫ＃　＝ｊ
409
リモートシステムからの応答待ち：書込完報告あり？
注）回復コピー中の更新Ｉ／Ｏに対する処理は省略
Y
本ペア（ｉ）に差分あり？
410
N
411
ペアｉ：回復到達報告（対リモート側）
412
回復コピー実行待ち（ペア１組を実行ペアに追加）
413
回復コピー完了報告（対リモート側）
終　了

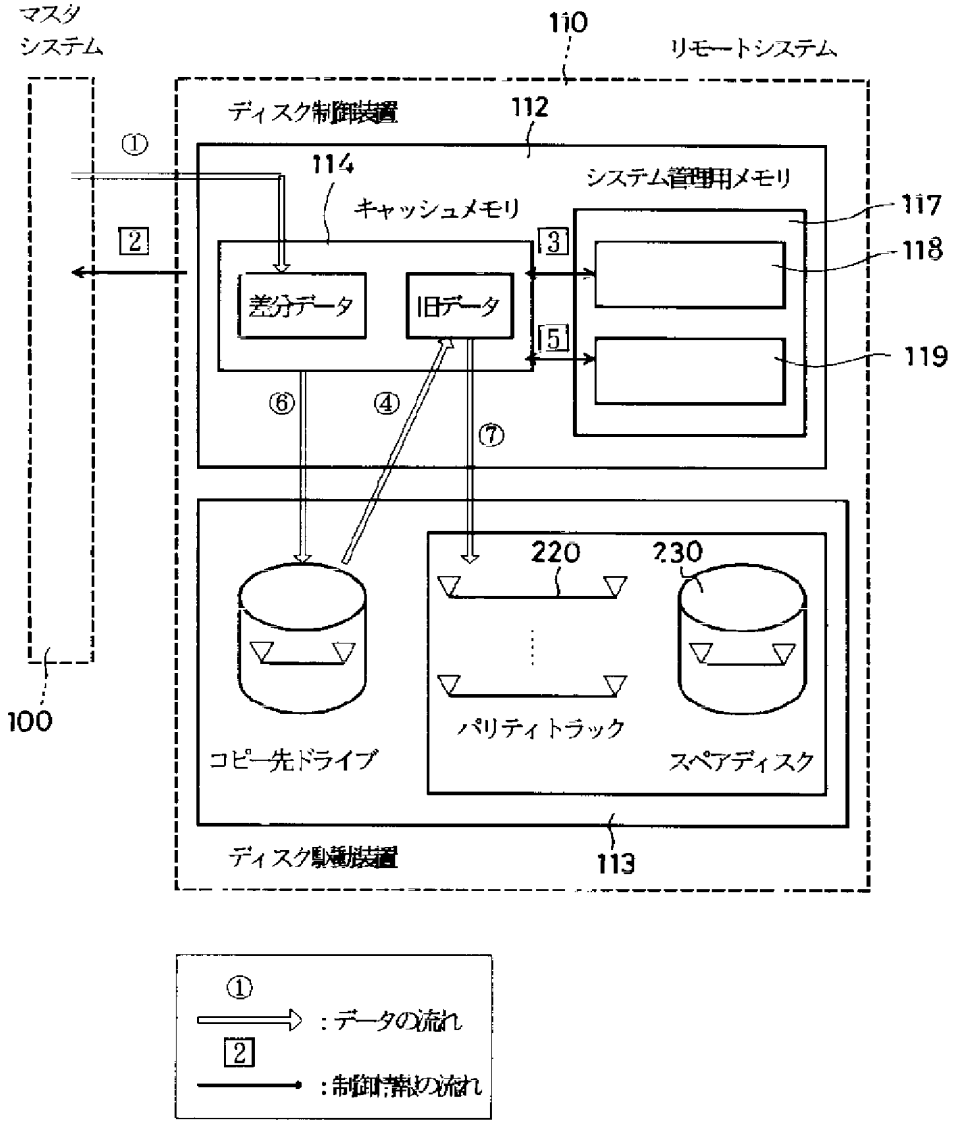
【図６】

図６


マスタシステム側から
回復コピー開始指示
回復コピー実行許容時間Tの送信
LDEV#=Ri，TRK#=jへの書込命令
回復コピー完了報告
リモート側の回復コピー処理
開　始
501
回復コピー指示あり？
502
回復コピー前処理
・T値の記憶、完了監視タイマスタート、etc
503
"回復コピー中"＝ON
マスタシステムへの応答（準備完了応答）
504
注）マスタ側の更新I／Oに対する更新コピーに対する処理は省略
注）516～519の処理は回復到達報告のなかったペアのみ実施する
505
"回復コピー中"＝ON？&トラックデータ書込有り？
N
Y
506
バックアップドライブへの格納アドレスを求める
507
キャッシュメモリ上に差分トラックデータを格納
508
差分トラックデータ書込終了報告
509
退避用の空きトラック検索
510
上書きされる旧データの退避（キャッシュメモリへのステージング）
N
使用可能キャッシュメモリ許容量オーバー？
511
Y
512
退避データの退避トラックへの書込み
514
回復コピー完了報告あり？
N
515
実行監視タイマ＞T？
N
Y
520
Y
"回復コピー中"＝OFF
516
退避データあり？
N
Y
517
退避データはキャッシュ上にあり？
N
Y
518
退避データの読み込み
519
バックアップドライブへの旧データ書込
終　了
注）"回復コピー中"＝ON　&回復到達報告のなかったペアは整合性回復処理が走ったことを表す。

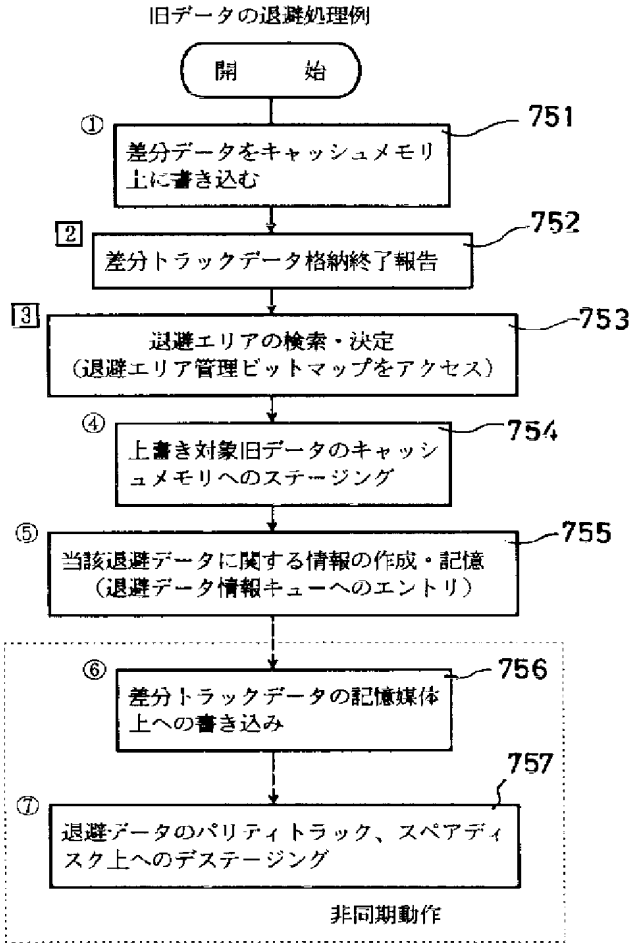
【図８】

図8


マスタ
システム
100
110
リモートシステム
ディスク制御装置
112
①
2
114
キャッシュメモリ
システム管理用メモリ
117
118
119
3
5
差分データ
旧データ
⑥
④
⑦
220
230
パリティトラック
コピー先ドライブ
スペアディスク
ディスク駆動装置
113
①
：データの流れ
2
：制御信号の流れ

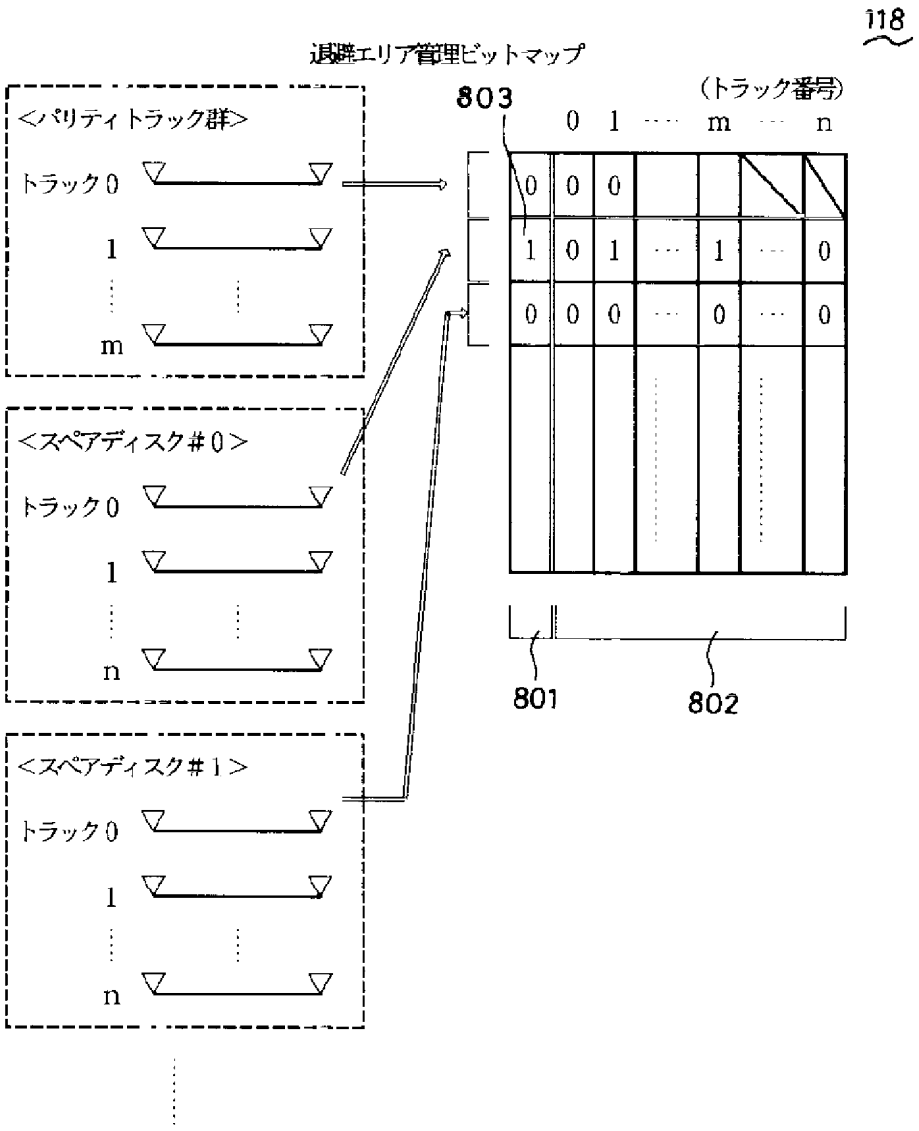
【図９】

図 9


旧データの退避処理例
開　始
①
差分データをキャッシュメモリ上に書き込む
751
2
差分トラックデータ格納終了報告
752
3
退避エリアの検索・決定
（退避エリア管理ビットマップをアクセス）
753
④
上書き対象旧データのキャッシュメモリへのステージング
754
⑤
当該退避データに関する情報の作成・記憶
（退避データ情報キューへのエントリ）
755
⑥
差分トラックデータの記憶媒体上への書き込み
756
⑦
退避データのパリティトラック、スペアディスク上へのデステージング
757
非同期動作

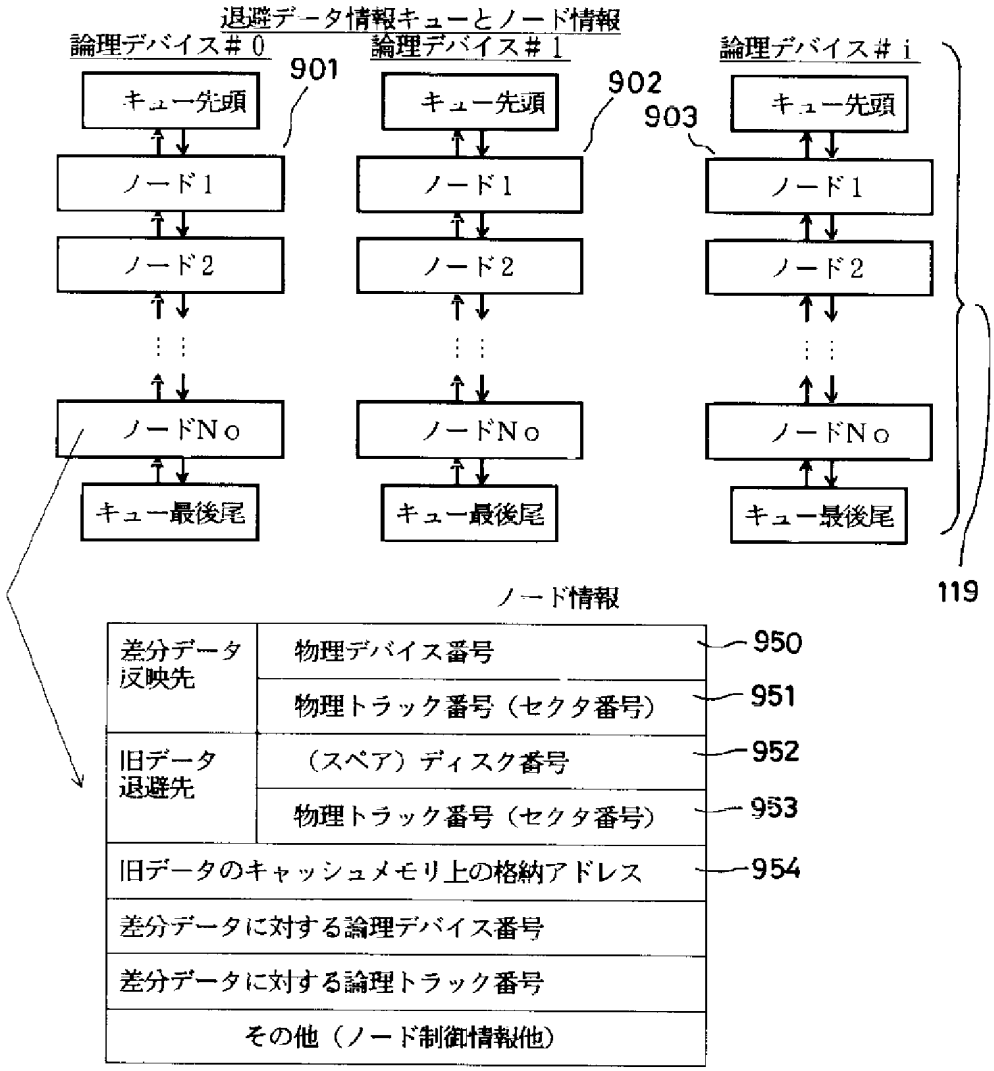
【図１０】

図１０


118
退避エリア管理ビットマップ
803
（トラック番号）
0 1 ・・・・ m ・・・ n
<パリティトラック群>
トラック0
1
m
<スペアディスク#0>
トラック0
1
n
<スペアディスク#1>
トラック0
1
n
0 0 0
1 0 1 ・・・ 1 ・・・ 0
0 0 0 ・・・ 0 ・・・ 0
801
802

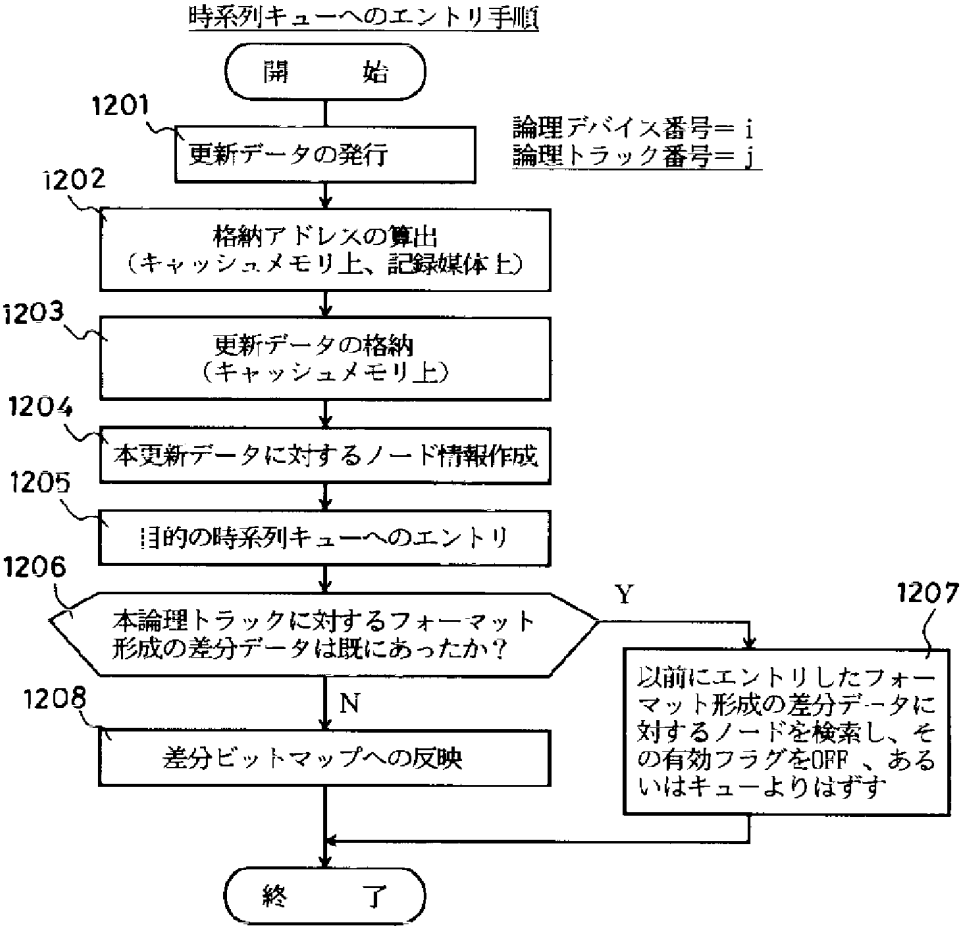
【図１１】

図１１


退避データ情報キューとノード情報
論理デバイス＃０
論理デバイス＃１
論理デバイス＃ｉ
901
902
903
キュー先頭
ノード１
ノード２
ノードＮｏ
キュー最後尾
119
ノード情報
差分データ反映先
物理デバイス番号
物理トラック番号（セクタ番号）
旧データ退避先
（スペア）ディスク番号
物理トラック番号（セクタ番号）
旧データのキャッシュメモリ上の格納アドレス
差分データに対する論理デバイス番号
差分データに対する論理トラック番号
その他（ノード制御情報他）
950
951
952
953
954

【図１３】

図１３


時系列キューへのエントリ手順
開　始
1201
更新データの発行
論理デバイス番号＝ｉ
論理トラック番号＝ｊ
1202
格納アドレスの算出
（キャッシュメモリ上、記録媒体上）
1203
更新データの格納
（キャッシュメモリ上）
1204
本更新データに対するノード情報作成
1205
目的の時系列キューへのエントリ
1206
本論理トラックに対するフォーマット
形成の差分データは既にあったか？
Y
1207
以前にエントリしたフォーマット形成の差分データに対するノードを検索し、その有効フラグをOFF 、あるいはキューよりはずす
N
1208
差分ビットマップへの反映
終　了